Control Engineering Practice 145 (2024) 105864

Contents lists available at ScienceDirect

Control Engineering Practice

journal homepage: www.elsevier.com/locate/conengprac

Automatic target recognition and geolocalisation of natural gas seeps using an autonomous underwater vehicle

Matteo Bresciani [a,b], Leonardo Zacchini [b,c], Alberto Topini [b,c], Alessandro Ridolfi [b,c], Riccardo Costanzi [a,b,d,*]

[a] Department of Information Engineering, Università di Pisa, via Girolamo Caruso 16, Pisa, 56122, Italy
[b] Interuniversity Center of Integrated Systems for the Marine Environment (ISME), Italy
[c] Department of Industrial Engineering, University of Florence, Florence, 50139, Italy
[d] Centro di Ricerca "E. Piaggio", University of Pisa, Largo Lucio Lazzarino 1, Pisa, 56122, Italy

ARTICLE INFO

Keywords:
Acoustic navigation and tracking
Cooperative ASV/AUV system
Navigation of autonomous marine robots
Automatic target recognition of seeps
Forward looking sonar for inspection
Convolutional neural networks

ABSTRACT

Autonomous Underwater Vehicles (AUVs) have proven to be a precious resource for oceans preservation thanks to their ability to accomplish survey missions on wide areas without the need of human supervision. This also includes the detection and mapping of underwater gas emissions, whether these are due to damaged offshore structures or naturally released from the seafloor. Indeed, submarine seepage have severe repercussions on the surrounding marine habitat and contribute to greenhouse gas emissions, hence the need to identify and monitor them. This work focuses on the online Automatic Target Recognition (ATR) and geolocalisation of underwater gas leakages utilising an AUV equipped with a Forward Looking Sonar (FLS). With the aim to accurately position the seeps, it was necessary to address the problem of navigating the vehicle in areas with the presence of gas bubbles. In fact, bubbles adversely affect acoustic sensors typically exploited by underwater robots for navigation. The paper investigates the effects that seepage have on the performance of navigation solutions based on two different acoustic sensors: a Doppler Velocity Log (DVL) sensor and an Ultra-Short BaseLine (USBL) device. Afterwards, a solution to autonomously recognise in real-time gas seeps on FLS acoustic imagery utilising a Single Shot MultiBox Detector (SSD) is presented. The relative position of the detected seep with respect to the vehicle is then retrieved and combined with the estimated AUV position to obtain the geodetic location of the seep. Finally, the proposed algorithm was tested during at sea experiments, where gas leaks were artificially reproduced, and the achieved results prove the validity of the proposed method to autonomously detect and accurately geolocate underwater seeps in real time using an AUV.

1. Introduction

Nowadays, Autonomous Underwater Vehicles (AUVs) are taking an increasingly central role in preservation and protection of marine environments (Bayat, Crasta, Crespi, Pascoal, & Ijspeert, 2017). In fact, recent developments in these technologies enable them to autonomously carry out missions to monitor water quality parameters (Amran et al., 2021), to assess the health state of marine life (Martin-Abadal, Guerrero-Font, Bonin-Font, & Gonzalez-Cid, 2018; Modasshir, Rahman, Youngquist, & Rekleitis, 2018), or to recognise possible harmful elements for the environment (Asada et al., 2010; Hwang, Bose, Nguyen, & Williams, 2020). Also included within the context of ocean conservation is the detection and mapping of marine gas seeps, mainly due to damaged man-made structures for oil and gas industry, natural seepage of methane and CO_2 from the seabed, or leaks related to Carbon Capture and Storage (CCS) process. Indeed, underwater gas seeps have a grievous impact on the surrounding marine environment (Jones et al., 2015; Taylor et al., 2015), and contribute to the amount of greenhouse gas released in the atmosphere (Blackford, Stahl, Bull, et al., 2014; Judd, 2003; Leifer, Kamerling, Luyendyk, & Wilson, 2010). Hence, their identification and localisation is essential to periodically monitor and quantify the amount of gas released, or to promptly intervene to repair a damaged structure.

This work proposes an online solution for autonomous recognition and accurate geolocation of natural gas seeps by utilising an AUV equipped with Forward Looking Sonar (FLS). To achieve such result, three basic steps were implemented: (i) accurately estimate the position

* Correspondence to: Dipartimento di Ingegneria dell'Informazione, Università di Pisa, via Girolamo Caruso 16, 56122, Pisa, Italy.
E-mail addresses: matteo.bresciani@phd.unipi.it (M. Bresciani), leonardo.zacchini@unifi.it (L. Zacchini), alberto.topini@unifi.it (A. Topini), alessandro.ridolfi@unifi.it (A. Ridolfi), riccardo.costanzi@unipi.it (R. Costanzi).

https://doi.org/10.1016/j.conengprac.2024.105864
Received 23 February 2023; Received in revised form 23 December 2023; Accepted 17 January 2024
Available online 28 January 2024

of the robot; (ii) automatically identify the gas seepage on sonar imagery; (iii) compute the geodetic coordinates of the leakage combining the information obtained in the previous steps.

The paper firstly addresses the problem of navigating with an underwater robot in areas with presence of gas bubbles, since an accurate estimation of the vehicle's position is a necessary (even if not sufficient) condition to obtain the correct location of a seepage. The consequences that gas leaks may have on the navigation of AUVs is a topic not extensively addressed in the literature; but it is reasonable to expect that interaction of acoustic waves with gas bubbles, and the associated environmental noise, negatively affect the acoustic sensors typically exploited by AUVs for navigation, resulting in a degradation in vehicle position estimation. Thus, the impact of bubbles on two different acoustic sensors, a Doppler Velocity Log (DVL) (Liu, Wang, Deng, & Fu, 2018) and an Ultra-Short BaseLine (USBL) device (Kebkal & Mashoshin, 2017), was studied, and the performance of tailored navigation strategies were evaluated. It is worth noting that the considered devices work on different frequency ranges: signals in the order of MHz for the DVL, and in the tens of kHz for the USBL. While the DVL-based navigation strategy consisted of traditional dead reckoning, the USBL-based solution made use of an Autonomous Surface Vehicle (ASV), which mounted the device and operated in support of the underwater robot's navigation, thus acting as a cooperative system. Finally, a third approach involving both sensors was also examined.

Afterwards, the manuscript introduces the Automatic Target Recognition (ATR) algorithm based on a Single Shot MultiBox Detector (SSD). The presented solution utilises FLS acoustic imagery to detect gas leaks in real-time, providing the identified object class, the classification confidence and the object bounding box coordinates in the image reference frame. Lastly, details are given on how to transform the bounding box centre from 2D pixel coordinates into the 3D vector containing the relative position of the detected leakage with respect to the AUV expressed in a North, East and Down (NED) reference frame. Combining this information with the absolute position estimated by the navigation strategy finally yields the geodetic coordinates of the seepage.

To summarise, the contribution of this work is twofold:

- to investigate the effects of gas bubbles on the navigation system of an underwater vehicle to enhance its localisation accuracy;
- to autonomously identify the position of a seepage in FLS images using a real-time CNN-based solution.

The proposed solution was tested during at sea experiments in which CO_2 leakages were artificially reproduced. The trials took place during April 2021 at the SEALab (La Spezia, Italy), the joint research laboratory between the Naval Support and Experimentation Centre of Italian Navy (CSSN) and the Interuniversity Centre of Integrated System for the Marine Environment (ISME). The data acquisition campaign was part of the "Detection And Mapping Of Submarine Seeps, Call 2" (DAMOSS2) activity, a TransNational Access (TNA) proposal within the EUMarineRobots (EUMR) H2020 European project framework. The TNA involved ISME, with the nodes of University of Pisa and University of Florence, as hosting partner and the Jacobs University of Bremen as Principal Investigator (PI). DAMOSS2 activity, built on top of the previous DAMOSS (Call 1) (Antonelli et al., 2021) and LAUA (Bresciani, Peralta, Ruscio, Costanzi et al., 2021) TNAs, had as goal to detect and accurately geolocalise CO_2 seeps exploiting heterogeneous exteroceptive sensors. A preliminary analysis of the impact of gas seepage on acoustic navigation systems used by AUVs has already been in Bresciani et al. (2022). However, that study was limited to a subset of only three missions and did not take into account the third navigation approach using information from both DVL and USBL sensors, which is instead analysed in this paper. Lastly, the deep learning approach for autonomously recognise the gas seeps on FLS images, as well as its integration with the navigation system to provide accurate geodetic coordinates of the leaks and the detailed analysis of the sensitivity of the performance of the three navigation solutions to the presence of seepage based on multiple experimental missions, represent original contributions of this manuscript.

The analysis of the navigation strategies, based on the acquired data, showed that both DVL and USBL sensors suffer from the presence of gas bubbles. In particular, the number of misleading readings produced by the DVL is considerably higher when the AUV moves on the seeps, leading to an increase in the drift of the error associated to the dead reckoning solution. Acoustic signals transmitted by the USBL, instead, suffer a performance degradation in terms of positioning rate and a lost of connectivity between the two robots. In contrast, the navigation solution involving both DVL and USBL measurements proved to be less sensitive to the disturbances caused by the seepage. The experimental trails also allowed to validate the developed CNN-based ATR methodology, which was running in real-time on board the AUV by means of a compact payload computer. The effectiveness of the method to autonomously detect CO_2 seepage on FLS acoustic images was demonstrated by the results, as well as its ability to obtain geolocated mapping of recognised leaks.

The manuscript is organised as follows: Section 2 presents a review of the literature for the methodologies typically used for seepage detection, the use of CNNs for target recognition, and the underwater navigation solutions. Section 3 describes the investigated navigation approaches based on DVL and USBL, while the developed ATR solution based on FLS acoustic imagery is detailed in Section 4. Following, a comprehensive description of the experimental activities, in which the solutions were tested, is provided in Section 5. Results are reported in Section 6 and discussed in Section 7. Finally, Section 8 summarises conclusions and possible future works.

2. Background

2.1. Methodologies for seeps detection

In the past few decades, sonars have proven to be a cost-effective solution to detect gas leaking from the seafloor, thanks to both their wide area coverage and the distinct difference in acoustic impedance between water and gas bubbles, which makes the latter clearly distinguishable (Carstensen & Foldy, 1947; Lurton, 2002). In the literature, several works tackle the problem of detecting seepage utilising different types of sonar systems: Nikolovska, Sahling, and Bohrmann (2008), Schneider von Deimling, Brockhoff, and Greinert (2007) and Urban, Köser, and Greinert (2017) exploit multibeam echosounders; Side Scan Sonars (SSS) are used in Klaucke, Weinrebe, Sahling, Bohrmann, and Börk (2005) and De Beukelaer, MacDonald, Guinnasso, and Murray (2003); while Weber, Jerram, and Mayer (2012) makes use of a split beam echosounder; lastly, the authors in Nadimi, Javidan, and Layeghi (2021) propose a solution based on Synthetic Aperture Sonar (SAS). These works consider the sonar systems mounted on the hulls of ships and offline data processing, with the leaks identification procedure executed manually. On the contrary, the authors in Blomberg, Sæbø, Hansen, Pedersen, and Austeng (2017) present a method for automatically detecting seepage using signal processing techniques applied to the data acquired by an interferometric SSS installed on an AUV. Furthermore, in Thorsnes, Chand, Brunstad, Lepland, and Lågstad (2019) a strategy to map cold seep habitats involving both ship and AUV is presented. The proposed approach consists of three phases, each of which exploits a different type of sensor: firstly water column data are collected using a multibeam echosounder mounted on the ship and used to detect gas flare; secondly an inspection with an AUV equipped with a long range sonar system (an interferometric SAS) is carried out to identify possible objects of interest; lastly, more detailed data are acquired with a camera system from areas of interest to confirm the presence of bubble streams. The solution of employing various type of sensors installed on different assets proves effective to perform surveys over very wide areas. Still, this strategy was not designed

to execute fully autonomous missions with an underwater robot, as seepage detection was not performed in real-time and the acquisition of new detailed data was done at a later time at known locations.

Alternative approaches to sonars make use of environmental probes to monitor the physical and chemical conditions of the water, as in Kato et al. (2017), Maeda, Shitashima, and Sakamoto (2015) and Monk et al. (2021). In fact, gas leaks cause water acidification resulting in a decrease in pH, which can be used as an indicator for detection. Despite the ability to identify seepage with these sensors was demonstrated, the methods are unable to accurately geolocate them.

The authors in Bhattarai et al. (2021), instead, investigate the use of an AUV provided with a camera system to identify and count bubbles using a deep learning approach. Although the results are impressive, the method is heavily dependent on visibility conditions, and it is thus unsuitable for uses in areas with high turbidity. Moreover, the presented image analysis was performed in post-processing, and the real-time application of the solution remains a challenge.

In this work, a FLS sensor mounted on an AUV was exploited to recognise in real-time the plume of gas bubbles released by submarine seepage. The choice to use a sonar sensor was dictated both by the fact that it is not affected by water turbidity and lighting conditions, and by its aforementioned advantages in terms of coverage and identifiability of leaks. Among the various sonar technologies, the FLS sensor was chosen as it can provide a visual representation of the insonified scene at each instant of acquisition, enabling the immediate use of image-based Deep Learning (DL) approaches without requiring further processing. Although other sonar systems such as SSS and SAS also provide acoustic images of the seafloor, their acquisition method involves an intrinsic delay and additional data processing to form a processable image. Moreover, these sensors impose limitations on the types of trajectories that the AUV must follow to avoid the occurrence of artefacts in the final image. FLS, on the other hand, does not impose any limitations on the motion of the robot, but rather allows for the implementation of obstacle avoidance strategies that could prove crucial to the safety of the AUV, especially during missions to inspect damaged anthropogenic structures or to monitor CO_2 seepage in volcanic areas characterised by shallow water — as in the case of DAMOSS1 TNA (Antonelli et al., 2021). Therefore, it is the authors' opinion that this sensor is more advantageous than others in the search for underwater seeps.

2.2. CNNs for automatic target recognition

CNNs have become the golden standard in many visual tasks, such as Image Classification and Object Detection (Krizhevsky, Sutskever, & Hinton, 2012). Indeed, their structure is optimised for analysing images and allows them to take advantage of some essential properties. Firstly, convolutional layers can extract relevant image features by learning the filter weights, whereas by adding pooling layers, more complex features can be extracted. In fact, the use of subsequent convolutional and pooling layers leads to a feature hierarchy, where simple features, such as edges and corners, are extracted in the first layers (Zeiler & Fergus, 2014), while more complex elements are identified in the last layers. Additionally, CNNs use computationally efficient filters that do not depend on the spatial position in the image and take into account that closer pixels are more likely to be correlated.

For what concerns the use of CNNs for Object Detection, several breakthrough architectures were proposed, e.g. You Only Look Once (YOLO) (Redmon, Divvala, Girshick, & Farhadi, 2016), SSD (Liu et al., 2016), and Faster R-CNN (Ren, He, Girshick, & Sun, 2015). These networks use different approaches but share a common working principle: the first layers compose the backbone that extracts the relevant features (thus, the backbone is also called 'feature extractor'), and then the final layers classify and localise possible objects of interest. Such image-based object detectors (in the following 'object detectors') take as input an image and give as output the predicted classes, together with the classification confidence and the coordinates of the bounding boxes containing the detected objects. Such coordinates are provided as the top-left and the bottom-right corners in the image reference frame, whose centre is in the image top-left corner and the x- and y-axis along the image width and height, respectively. For the sake of the clarity, a general object detector workflow is depicted in Fig. 1.

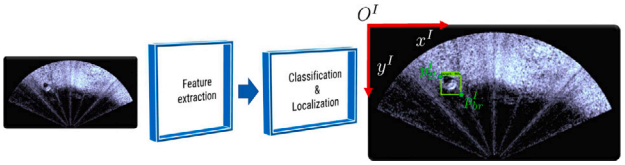


Fig. 1. A general workflow of an object detector. Given an image as input, it outputs the predicted classes, with the computed confidence, and the object bounding boxes coordinates, expressed as the top-left (p^I_{tl}) and the bottom-right (p^I_{br}) corners in the image reference frame $\{O^I x^I y^I\}$.

The marine community has investigated the use of CNNs for tackling the object detection task, also called ATR, on acoustic images. In fact, in recent years, CNNs have shown better performance than template-matching-based approaches (Jin, Liang, & Yang, 2019; Valdenegro-Toro, 2016) that cannot generalise template patterns and suffers from handling multi-scale objects. Nevertheless, developing a high-performing CNN architecture from scratch is by no means straightforward and is time and cost expensive (Jin et al., 2019; Palomeras, Furfaro, Williams, Carreras, & Dugelay, 2021). In addition, the limited space usually available on compact AUVs imposes the use of small payload hardware with limited computational capabilities, posing additional challenges in developing CNN solutions for real-time applications. Hence, the application of state-of-the-art object detection CNNs for analysing acoustic images emerged as a simple but extremely effective way to go. The usage of the YOLO network allowed Kvasic and colleagues (Kvasić, Mišković, & Vukić, 2019) to develop a detection and tracking solution for a robot-diver collaborative system. In Kim and Yu (2016), YOLO was utilised for developing a real-time object-detection algorithm for localising an agent vehicle with respect to the main AUV for manipulation tasks. Then, in Zacchini, Ridolfi et al. (2020) and Zacchini, Franchi et al. (2020), SSD and the Faster R-CNN architectures were employed to identify Objects of Potential Interest (OPIs) in FLS images. Although, as described in Section 2.1, the literature presents some methodologies to autonomously recognise and geolocalise natural gas seeps (i.e. the goal of this work), solutions leveraging CNN architectures applied on FLS imagery are still an open point investigated in this paper.

2.3. Navigation in the underwater domain

The navigation of an autonomous vehicle is, in general, a nontrivial problem and it becomes even more complex when the robot has to move in a challenging environment, as the underwater domain is. In fact, the marine domain presents several technological challenges mainly related to the electromagnetic signals attenuation in water. As consequence, the GPS data typically exploited to determine the robot position in terrestrial applications it are not available underwater, thus the need for alternative localisation systems arises. Paull, Saeedi, Seto, and Li (2014) presents a comprehensive review of the state-of-the-art navigation and localisation methods for AUVs. While Maurelli, Krupiński, Xiang, and Petillot (2022) focuses its review on both passive and active techniques whose effectiveness was proven in the field. Regarding this work, the effects that gas bubbles have on the navigation performance of underwater vehicles is, to the best of the authors' knowledge, a topic not covered in literature. Nonetheless, it is reasonable to expect consequences on acoustic sensors used in the navigation of AUVs that cannot be neglected when planning an inspection in areas with bubbles streams presence, since it may compromise the navigation of the vehicle and therefore its ability to safely complete the assigned task.

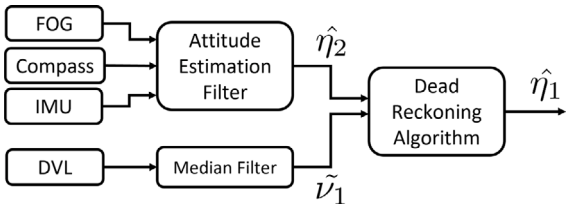


Fig. 2. Block diagram of the DVL-based approach.

3. Navigation strategies

This section presents three acoustic sensor-based navigation strategies considered to evaluate the influence of seepage on underwater vehicle navigation.

3.1. DVL-based approach

First, a dead reckoning method relying on a DVL sensor is considered. Similar algorithms are often employed on mobile robots equipped with reliable sensors that provide high-precision measurements, such as DVL and Fibre Optic Gyroscope (FOG). The approach requires the knowledge of the vehicle velocity and attitude, which are then integrated over a fixed step time to obtain the position. While the velocity of the robot is provided by the DVL readings, appropriately filtered through a median filter with a moving window of 5 samples to mitigate the effects of outliers, the attitude is estimated by a non-linear observer. Specifically, the observer produces an attitude estimate $\hat{\boldsymbol{\eta}}_2$ fusing MEMS Inertial Measurement Unit (IMU), compass, and single-axis FOG data, as described in Costanzi, Fanelli, Monni, Ridolfi, and Allotta (2016). Then, the attitude estimate is combined with the DVL filtered data $\tilde{\boldsymbol{\nu}}_1$ to estimate the AUV position $\hat{\boldsymbol{\eta}}_1$, as shown by the block diagram in Fig. 2. It is worth mentioning that the DVL utilised within this work has four acoustic beams oriented in a diverging configuration and it operates with acoustic signals at a frequency of 1 MHz. Once the device detects the bottom with at least three of the four beams, it calculates the velocity along each beam and provides a velocity measurement in a XYZ sensor reference frame (Nortek Manuals, 2022), otherwise the instruments readings are flagged as invalid.

3.2. USBL-based approach

The second investigated approach makes use of an USBL device and leverages on ASV/AUV cooperation to estimate the AUV position. The strategy, presented by the authors in Bresciani, Peralta, Ruscio, Bazzarello et al. (2021), requires to install an USBL device with communication capabilities on an ASV, while the AUV must be equipped with a compatible acoustic modem to receive the data on its position that were retrieved by the ASV. So, thanks to a specifically designed acoustic protocol, the ASV is able to acquire the position of the underwater robot without requiring any prior synchronisation between the vehicles' clocks and prioritising the positioning rate. In particular, the relative position of the AUV with respect to the ASV is measured by the USBL and compensated with the attitude, provided by the Attitude and Heading Reference System (AHRS) integrated within the USBL, and the GPS mounted on the surface vehicle. This allows to obtain the geodetic coordinates of the AUV, which are then transmitted to the latter thanks to the acoustic protocol. Eventually, the received measurements are utilised by the AUV in a navigation filter able to deal with delayed observations to compute the past corrected state estimate and then propagate the result up to the current time (Costanzi et al., 2018). The filter implemented within this work is an Extended Kalman Filter (EKF) with a kinematic model and a state vector expressed as $\mathbf{x} = [n, e, s, \delta]^\intercal \in \mathbb{R}^4$, according to the *SNAME* convention (Fossen, 2002). In details, n and e represent the vehicle position with respect to a NED reference frame $\{n\}$, s represents its speed and δ the drift angle, described as the difference between the course and heading angles. Hence, the discrete time state transition model can be defined as:

$$\begin{cases} n_{k+1} = n_k + dT\cos(\psi_k + w_\psi + \delta_k)s_k \\ e_{k+1} = e_k + dT\sin(\psi_k + w_\psi + \delta_k)s_k \\ s_{k+1} = s_k + dT w_s \\ \delta_{k+1} = \delta_k + dT w_\delta \end{cases} \tag{1}$$

In these equations, ψ represents the yaw angle in the $\hat{\eta}_2$ estimate, obtained by the non-linear observer described in Section 3.1, and here exploited as input to the model. In addition, dT is the discretisation time, while w_ψ, w_s, w_δ are white Gaussian process noises. The received AUV absolute positions are then expressed in the NED frame and, if considered valid by an outliers rejection procedure based on Mahalanobis distance, are used as corrections. The implemented observation model is the following:

$$\begin{cases} y_n = n + \lambda_p \\ y_e = e + \lambda_p \end{cases} \tag{2}$$

where λ_p is a white Gaussian additive noise, equal on both North and East directions. It should be noted that transmitting the geodetic coordinates to the underwater robots allows to completely decouple the ASV and AUV systems, since such data are not affected by the relative motion of the two vehicles. Regarding the USBL device employed in this work, instead, its operating frequency range is 18–34 kHz, which differs significantly from those used by the DVL. Furthermore, it is noteworthy that the depth of the underwater robot, acquired through an independent depth sensor integrated into the AUV, is combined with the 2D position estimated by the navigation filter to retrieve the 3D position of the vehicle with respect to the NED frame. Lastly, the authors would like to point out that an inverted-USBL configuration, as the one described in Morgado, Oliveira, Silvestre, and Vasconcelos (2007) and Rypkema, Fischell, and Schmidt (2017), would also be a viable alternative. Mounting the USBL device on the AUV would indeed facilitate the determination of the relative position between the underwater robot and the target beacon, thereby obviating the need for supplementary acoustic communication. However, in the specific scenario described in this study, the position information is inherently relative to the position of the ASV, which is subject to change. Consequently, an additional communication step would remain indispensable to convey the geodetic coordinates of the ASV to the underwater robot, so that the latter can calculate its absolute position and enhance its navigation capabilities.

3.3. Combined approach

A navigation strategy relying on both DVL and USBL sensors was also investigated. Similar sensor fusion solutions are widely employed in literature to exploit the full set of readings available to the underwater robot, as in Font, Bonin-Font, Negre, Massot, and Oliver (2017) and Rigby, Pizarro, and Williams (2006). Within this work, the aim of the approach was to assess whether a combination of the two technologies could overcome the negative influence that gas bubbles have on each instrument, and thus improve the accuracy of the estimation. For this purpose, this solution implements an EKF algorithm with a kinematic model akin to the one presented in Section 3.2, but with a different state definition to better exploits the additional information obtained by the DVL. Indeed, since the linear velocity components are provided to the vehicle control system expressed in the Body reference system $\{b\}$, already compensating for the DVL pose, the data can be used to retrieve both velocity and course angle information. The filter state can be then formalised as $\mathbf{x} = [\boldsymbol{\eta}_1, \boldsymbol{\nu}_1]^T \in \mathbb{R}^6$, where $\boldsymbol{\eta}_1 = [n, e, d]^T \in \mathbb{R}^3$ contains the position of the AUV along the axes of the NED reference frame and $\boldsymbol{\nu}_1 = [u, v, w]^T \in \mathbb{R}^3$ represents the body-fixed linear velocities with respect to the surge, sway and heave

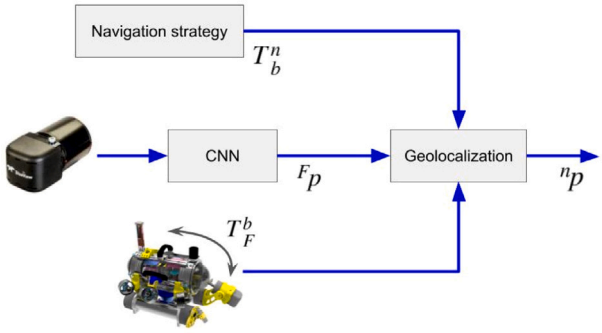


Fig. 3. Block diagram of the ATR method to identify and localise gas bubbles streams using FLS imagery.

axes. A slightly different model of state transition ensues, which can be described as:

$$\begin{cases} \boldsymbol{\eta}_{1k+1} = \boldsymbol{\eta}_{1k} + dT\mathbf{R}_{\mathbf{b}}^{\mathbf{n}}(\boldsymbol{\eta}_{2k})\boldsymbol{\nu}_{1k} \\ \boldsymbol{\nu}_{1k+1} = \boldsymbol{\nu}_{1k} + dT\mathbf{W}_k \end{cases} \tag{3}$$

where $\boldsymbol{\eta}_2 = [\phi, \theta, \psi]^T \in \mathbb{R}^3$ are the Euler angles representing the attitude of the vehicle and used to express the body-fixed velocities in the NED reference frame through the rotation matrix $\mathbf{R}_{\mathbf{b}}^{\mathbf{n}}(\boldsymbol{\eta}_{2k})$, whilst $\mathbf{W}_k \in \mathbb{R}^3$ represents a white Gaussian noise vector. Also for this method, the estimate $\hat{\boldsymbol{\eta}}_2$ computed by the nonlinear observer was considered as attitude $\boldsymbol{\eta}_2$ of the robot, while the DVL readings are used as direct observations of the state velocities, according to the following equation:

$$\mathbf{y}_{\boldsymbol{\nu}_1} = \boldsymbol{\nu}_1 + \zeta_{\boldsymbol{\nu}_1} \tag{4}$$

where $\zeta_{\boldsymbol{\nu}_1}$ is a white Gaussian additive noise vector. Furthermore, the depth sensor and the USBL device are used to correct the position estimate. The latter using the same observation model described in Eq. (2). Finally, the rate at which each sensor gives a new measurements, whether valid or invalid, differs considerably. Although for the depth sensor and DVL the rate is fixed, for the USBL device such value varies according to several factors, such as: interference in the acoustic channel, inter-vehicular distance, operational scenario, etc. Nevertheless, the EKF handles this discrepancy autonomously, applying the appropriate correction only when a measurement is available to the robot.

4. Automatic seeps recognition and localisation

When a new FLS image is acquired, it is analysed by the ATR software by means of the trained CNN model. As mentioned in Section 2.2, the network outputs the predicted object classes, with the computed confidence level and the object bounding boxes, which are then used to localise the detected gas seeps, according to the scheme shown in Fig. 3.

In the following, further details about the network selection, training and validation processes are discussed. Lastly, the procedure used to retrieve the geodetic coordinates of the identified seeps is described.

4.1. Network selection

To select a proper CNN architecture, several aspects shall be taken into account. The lack of an available dataset of FLS images and the aforementioned difficulties regarding the development of a CNN for real-time applications have pushed the focus of this research towards the use of state-of-the-art solutions on which apply a transfer learning procedure (Weiss, Khoshgoftaar, & Wang, 2016). Indeed, transfer learning techniques make it possible to use pre-trained CNNs on vast generic datasets, and then fine-tuning only the final layers to recognise a specific set of OPIs. These methods represents a good solution to overcome the lack of large dataset for the training, as in the case of acoustic images from FLS. Therefore, the network architecture consist of a 'feature extractor', a pre-trained backbone in charge of extracting the most relevant features, and an 'object detector', the final layers optimised to detect natural gas seeps in FLS images. The TensorFlow Object Detection API allows to easily exploit the aforementioned strategy to train modular CNNs on a custom dataset. As described in Huang et al. (2017), the accuracy, measured as the mean Average Precision (mAP), is not the only parameter that shall be considered when developing an ATR solution for real deployments, but also the inference speed and the memory demand play a fundamental role. To help in selecting only the most promising CNN architectures, Huang and colleagues analysed the performance of three meta-architectures (i.e. SSD, Faster R-CNN, and Region-based Fully Convolutional Network (R-FCN)) with six feature extractors. The speed/accuracy trade-off was investigated and the optimal frontier, i.e. where the best detectors can be found, was defined. The meta-architecture performance is strongly connected with the choice of the backbone, but in general terms, the Faster R-CNN is more accurate, while the SSD and R-FCN are faster. Finally, another aspect taken into consideration during network selection was the available hardware. Since the algorithm has to run on board an underwater vehicle, which generally has limited computational capabilities, and based on the authors' previous experience (Zacchini, Franchi et al., 2020; Zacchini, Ridolfi et al., 2020; Zacchini et al., 2022), the focus shifted to a SSD network with MobileNet as backbone (SSD MobileNet v2). Indeed, the chosen SSD network, along with the derived ATR system based on FLS imagery, underwent thorough validation across diverse scenarios, including anti-submarine warfare and underwater inspection of man-made structures, as detailed in the referenced works. This comprehensive validation process has led the authors to the informed conclusion that the selected network represents an optimal compromise that meets the various aforementioned requirements, and is therefore promising for potential applications in other contexts, including gas seepage detection.

4.2. Network training & validation

In the context of the previous DAMOSS TNA (Call 1), an heterogeneous dataset consisting of 128 FLS images of artificially reproduced submarine gas seeps was acquired (Bucci, Zacchini, Franchi, Ridolfi, & Allotta, 2022). In particular, the sonar employed in this research had a field of view of 130° and provided images with a resolution 894 × 477 pixels. The entire dataset was then divided as follows: 108 images were used for training, while the other 20 images were used as test set, enabling the network evaluation to avoid overfitting issues. In detail, the training was performed on a laptop fitted with 16 GB RAM, an Intel Core i7-8750H processor, and an Nvidia GeForce GTX 1050 Ti card. The SSD network was trained using the RMSProp algorithm as the optimiser and a batch size of 24. The network was constantly evaluated during the training that lasted 10k steps, as shown in Fig. 4. After about 1.8k steps of training, a clear overfitting behaviour can be noted. Consequently, the network weights corresponding to the minimum of the loss were utilised during the experimental tests.

4.3. Seeps geolocalisation

Once the leakage has been recognised on the FLS image by the proposed ATR solution, it needs to be geolocalised. To this end, the seep position, expressed as 2D pixel coordinates in the image reference frame $\langle I \rangle$, is firstly transformed into a 3D position vector with respect to the reference frame $\{F\}$ rigidly attached to the FLS sensor, with the x-axis points forward and the z-axis points downwards. Thereafter, the data is expressed in a NED reference frame and combined with the AUV absolute position to obtain the seep geodetic coordinates. In detail, the CNN solutions outputs a bounding box for each detected object of interest expressed as the top-left and the bottom-right corners in the

Fig. 4. The loss curve obtained by evaluating the SSD model with the validation dataset.

image reference frame $\langle I \rangle$. The bounding boxes are predicted so that the detected object is contained within them, thus it is reasonable to select the centres of the boxes as reference points on which to perform the geolocalisation of the objects. However, due to the loss of the information about the elevation angle in the 3D to 2D image formation process, only the range $\bar{R}$ and the azimuth angle α can be calculated from an FLS image (Franchi, Ridolfi, & Allotta, 2021; Negahdaripour, 2013). Given a point on an FLS image $p^I = [x \ y]^T$, its 2D position in spherical coordinates (natively used in the imaging process) with respect to the FLS can be computed as:

$$\begin{bmatrix} \bar{R} \\ \alpha \end{bmatrix} = \begin{bmatrix} \tilde{d}\sqrt{x^2+y^2} \\ atan2(y,x) \end{bmatrix}, \tag{5}$$

where $\tilde{d}$ denotes the ratio between the FLS range and the image's height, which is used to convert the distance from pixels to meters. Furthermore, by using the FLS tilt angle with respect to the horizontal plane γ, an estimated 3D seep position in the $\{F\}$ frame, $p^F = [X, Y, Z]^T$, can be computed as:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \bar{R} \begin{bmatrix} \cos\gamma\cos\alpha \\ \cos\gamma\sin\alpha \\ -\sin\gamma \end{bmatrix}. \tag{6}$$

Then, by knowing the FLS mounting pose, it is possible to express the position of the seeps in the body reference frame $\{b\}$. Introducing the transformation matrix $\boldsymbol{T}$ of the special Euclidean group in $\mathbb{R}^3$:

$$\mathrm{SE}(3) := \left\{ \boldsymbol{T} = \begin{bmatrix} R & \boldsymbol{p} \\ \boldsymbol{0}^T & 1 \end{bmatrix} \mid R \in \mathrm{SO}(3), \boldsymbol{p} \in \mathbb{R}^3 \right\}, \tag{7}$$

and the four-dimensional homogeneous representation :

$$\widetilde{p} = \begin{bmatrix} p \\ 1 \end{bmatrix}, \tag{8}$$

the estimated seep position with respect to the body frame $\{b\}$ can be expressed in a compact notation as:

$$\widetilde{\boldsymbol{p}}^b = \boldsymbol{T}_{\boldsymbol{F}}^{\boldsymbol{b}} \widetilde{\boldsymbol{p}}^{\boldsymbol{F}}. \tag{9}$$

It is worth highlight that in the above equations, $\boldsymbol{p}$ represents a position vector, R is a rotation matrix (and as such it is part of the special orthogonal group in $\mathbb{R}^3$), while $\boldsymbol{T}_{\boldsymbol{F}}^{\boldsymbol{b}}$ depends on the FLS mounting pose. Afterwards, the seeps position can be projected into the NED frame according to the following equation:

$$\widetilde{\boldsymbol{p}}^{\boldsymbol{n}} = \boldsymbol{T}_{\boldsymbol{b}}^{\boldsymbol{n}} \widetilde{\boldsymbol{p}}^{\boldsymbol{b}} = \boldsymbol{T}_{\boldsymbol{b}}^{\boldsymbol{n}} \boldsymbol{T}_{\boldsymbol{F}}^{\boldsymbol{b}} \widetilde{\boldsymbol{p}}^{\boldsymbol{F}}, \tag{10}$$

where the transformation matrix $\boldsymbol{T}_{\boldsymbol{b}}^{\boldsymbol{n}}$ is composed of the estimate of the AUV position expressed in NED frame $\hat{\eta}_1$, provided by the implemented navigation strategy, and the rotation matrix R_b^n, obtained from the estimate of the vehicle attitude $\hat{\eta}_2$. Finally, by applying a simple conversion to the position in NED frame $\widetilde{\boldsymbol{p}}^{\boldsymbol{n}}$ is possible to map the seep by retrieving its latitude and longitude coordinates.

Fig. 5. Satellite representation of the operational area, in green, while the position of the deployed pipe used to artificially reproduce the gas streams is represented in cyan.

Fig. 6. Example of gas streams artificially reproduced within the operational area during the experimental activity.

5. Experimental activities

Both navigation strategies and ATR solution were tested and validated during at sea experiments at the SEALab laboratory, in La Spezia (Italy). The operational area in which the tests took place is located within an artificial basin confined between two piers, where the depth varies from 3 m to 15 m, near the docks and in the middle of the basin, respectively. Therefore, the scenario results challenging from the acoustic point of view, since the very shallow waters may imply frequent multipath phenomena causing outliers in the USBL measurements or loss of communication. To perform the tests, an underwater gas seepage was artificially reproduced by drilling a rubber tube at regular intervals of 30 cm and placing it on the seabed by means of ballasts. Then, using a compressor, the tube was filled with pressurised air, producing columns of bubbles from each of its holes. The geodetic coordinates of the pipe were measured during deployment by the operators aboard a Rigid Hull Inflatable Boat (RHIB) using a portable GPS device. Fig. 5 reports a satellite view of the operational area, where the location of the deployed pipe is depicted in cyan. Furthermore, an example of the artificial gas leaks reproduced within the experiments is shown in Fig. 6.

5.1. Equipment setup

The vehicles involved in the experimental activities were the Feel-Hippo AUV, designed and developed by the Department of Industrial

Table 1
FeelHippo AUV sensor set.

Type	Model
GPS	U-blox NEO-7P
IMU	Xsens MTi-300
FOG	KVH DSP 1760
FLS	Teledyne Blueview M900 2D
DVL	Nortek DVL1000
Depth sensor	integrated within the DVL
Acoustic modem	EvoLogics S2CR 18/34

Table 2
MGB300 ASV sensor set.

Type	Model
GPS	Ublox Neo 6 m
AHRS	Orientus Advanced Navigation
Depth sensor	Keller PA20Y
USBL	EvoLogics S2CR 18/34

Engineering of the University of Florence (UNIFI DIEF), and a Mobile Gateway Buoy 300 (MGB300) ASV, one of the CSSN assets. FeelHippo is a lightweight and compact AUV capable to control four of the six DOF thanks to the disposition of its propulsion system (Ridolfi et al., 2016). In 2017, the robot underwent a major overhaul to enhance its performance both in terms of mechanical components and system autonomy, (Franchi, Fanelli, Bianchi, Ridolfi, & Allotta, 2020) and is now able to safely perform autonomous underwater missions. This is also due to its wide range of sensors: GPS, depth sensor, DVL, IMU, single-axis FOG, acoustic modem, and FLS. Table 1 reports a comprehensive view of the AUV sensors set and models, while Fig. 7(a) shows the robot during an underwater mission in a volcanic area with high presence of CO_2 seeps. It is worth mentioning that the acoustic modem is mounted on top of the vehicle hull, therefore the robot has to be completely submerged to exchange acoustic messages with any other compatible modem. This resulted in the inability to acquire a GPS ground truth during the experiments since FeelHippo had to be entirely underwater to be localised with the USBL device. Moreover, to run the developed CNN-based ATR solution, an NVIDIA Jetson Nano (NVIDIA, 2018) was mounted on the AUV. The NVIDIA Jetson Nano is a platform for embedded AI computing specifically designed to meet the power consumption and limited space requirements of autonomous machines. The platform acted as a dedicated payload computer that was connected to the vehicle main computer, in charge of running onboard processing navigation and control algorithms and supervising the state of the vehicle, through an Ethernet cable. Nevertheless, despite its small size allowed it to be installed into the limited space available on the underwater robot, its limited computational capabilities have imposed constraints on the network selection.

The MGB300 ASV, instead, is a torpedo-shape vehicle designed to support an AUV during its missions and to serve as multi-domain communication bridge between the Command and Control Station (C2S) and the underwater robot. To this end, the ASV mounts WiFi and radio antennas on an aluminum structure for better signal coverage, while an USBL device was installed below the vehicle hull. The Ultra High Frequency (UHF) radio modem allows the long range communication between the ASV and a C2S, while the WiFi connection (when available) enables the share of information with an higher data rate and reliability compared to radio and acoustic channels. MGB300 is also equipped with GPS and AHRS, used to measure its position and to estimate its orientation, respectively. Fig. 7(b) displays an image of the robot, while Table 2 indicates the vehicle's sensor suite, with details on the models. Lastly, Table 3 summarises the main characteristics of the DVL and USBL devices mounted on the vehicles and utilised within the experiments.

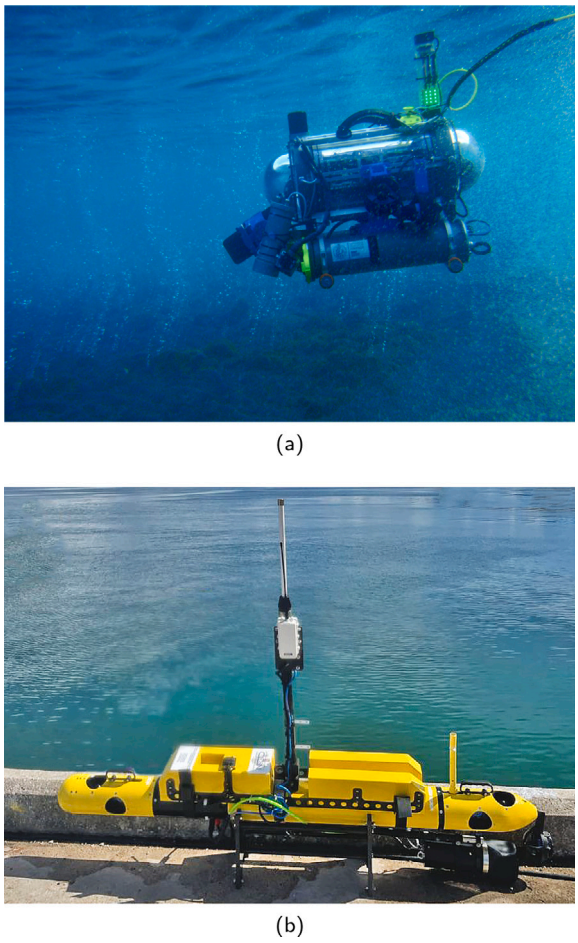


Fig. 7. AUV and ASV exploited in the experimental activities (a) FeelHippo AUV navigating underwater in a volcanic area with high presence of CO_2 seeps during at sea tests carried out within the DAMOSS1 project, in Vulcano (Italy) (b) MGB300 ASV on a stand before deployment at sea.

Table 3
DVL and USBL main specifications.

	DVL	USBL
Frequency band	0.75–1.25 [MHz]	18–34 [kHz]
Accuracy	±0.1 [cm/s]	0.01 [m]
Resolution	0.01 [mm/s]	0.1 [°]
Dimensions	ø114 × 158 [mm]	ø110 × 170 [mm]
Weight (dry/wet)	1.60/0.15 [kg]	5.78/0.73 [kg]

5.2. Experiments description

The experiments were designed to evaluate the performance of the navigation strategies and the ability of the CNN-based ATR method to identify and accurately geolocate the streams of gas bubbles. For this purpose, the AUV carried out predefined lawn-mower trajectories at 2 m depth and 0.2 $\frac{m}{s}$ speed, some of which were planned to straddle the pipe, and thus to cross the artificial seepage, while others were planned to inspect the gas columns while remaining on the same side with respect to the pipe. The choice of using this speed value was to limit the acoustic disturbance produced by the vehicle thrusters, thus minimising any possible repercussions on the acoustic sensors under study. The trajectories were executed by starting near the dock and moving towards the centre of the basin to reduce the risks of possible collisions of the vehicle with the pier itself. In addition, the AUV navigated by taking advantage of the dead reckoning algorithm described in Section 3.1, although the EKF estimates were computed

online but not utilised in the vehicle control loop, for both the USBL-based and combined approaches. The ASV, instead, was hovering in a fixed position during FeelHippo localisation in all but one trial (more detailed later). This was done both for safety reasons, thus avoiding possible collisions between vehicles, and to reduce the influence that the movement of the surface robot could have on the accuracy of the acoustic localisation system. Indeed, although in general having the ASV following the AUV can lead to improvements in localisation results due to a reduction the inter-vehicle distance, in the case presented this improvement would have been negligible compared to the risk of vehicle collision. This was due to the size of the operating area, which implied a limited inter-vehicle distance throughout the duration of the trials. A more in-depth analysis of the effects of inter-vehicle distance on localisation accuracy, however, is beyond the scope of this paper.

To study the impact of gas bubbles on the acoustic sensors and proposed navigation solutions, four configurations of the two vehicles with respect to the pipe were tested, each of which was designed with a specific objective. In particular, it was necessary to verify at first if the presence of gas leaks in the nearby (tens of meters) of the operational area had consequences on the exploited acoustic sensors (*S1*). This was important to evaluate the performance of the systems in a benchmark situation and, eventually, to use these results in a comparative analysis with other scenarios. Then, the authors wanted to investigate the effects of bubbles on the USBL device, so a configuration in which seeps are interposed between the ASV and AUV was considered (*S2*). In a similar scenario is indeed reasonable to assume that gas bubbles may interfere significantly with the positioning system, but it is not known to what extent. Next, the performance of the DVL sensor were examined in a scenario (*S3*) in which the AUV was moving above the gas leaks. In fact, because of the way the DVL is mounted (down-looking), it is reasonable to suppose that bubbles may have an effect on the sensor only when the vehicle passes over the pipe. Finally, a single test was conducted with MGB300 going back and forth in the middle of the basin (*S4*) to assess the performance of the positioning system during movement, and as a solution to avoid the acoustic obstruction generated by gas streams. Specifically, the surface vehicle was executing in loop an autonomous mission consisting of two waypoints placed in the centre of the basin. No coordination of movement between the two vehicles was implemented, resulting in unpredictable configurations of the vehicles with respect to artificial seepage. Operationally, the various configurations were implemented as follows:

(*S1*) AUV and ASV operating on the same side with respect to the pipe, in an area sufficiently distant from the seeps location.;
(*S2*) AUV moving as in (*S1*), but ASV located at the opposite side of the pipe;
(*S3*) ASV positioned on one side of the pipe and AUV straddling both sides;
(*S4*) AUV crossing the seepage and ASV moving back and forth in the basin.

Fig. 8 shows examples of the relative configurations between the two robots and the pipe that characterise the various scenarios described. Table 4, instead, summarises the entire dataset of 10 missions carried out during the two days of DAMOSS2 project experiments, grouped by reference scenario and listing for each test: an identification number (ID), the length of the reference trajectory (PL), and the number of DVL and USBL data collected. During the first day of experiments the missions with ID 1, 2, 3, 5 and 6 were performed, while missions 4, 7, 8, 9 and 10 were conducted in the second day.

The paths planned for FeelHippo AUV varied slightly among the different missions in the attempt to maintain the vehicle close enough to seeps so that they would remain within the operational range of the FLS. Indeed, the experiments had the twofold objective of evaluating navigation performance and validating the proposed ATR algorithm in real time and in the field. To this end, the AUV had to autonomously recognise gas seeps through FLS images and then provide their corresponding geodetic coordinates. Moreover, the paths had also to adapt to the characteristics of each scenario to preserve the safety of the robot within the operational area. However, during the second day of testing, an hardware problem in the underwater connector of the FLS occurred, making it impossible to collect data with the sensor. As result, the dataset acquired during the experimental campaign contains FLS images only for the missions executed during the first day of trials.

Table 4
Dataset of mission performed during the experimental trials.

ID	Scenario	PL [m]	# DVL data	# USBL data
1	*S1*	137	2857	791
2		79	1779	530
3		99	2224	548
4		88	2185	498
5	*S2*	115	2485	621
6		97	2130	624
7	*S3*	106	2485	550
8		113	2597	569
9		110	2371	576
10	*S4*	108	2728	601

6. Metrics & results

6.1. Navigation performance

Different metrics were computed to assess the impact of natural gas on the acoustic sensors employed in the work and to compare the proposed navigation solutions. For the DVL-based approach the percentage of invalid measurements (obtained as described in Section 3.1) was considered as indicator of bubbles effect on the sensor. For what it concerns the USBL-based solution, the examined parameters include: the percentage of lost packets (l_p), the maximum time interval without any new measurement Δ_t, and the average rate of USBL positioning $\overline{\chi}$. In addition, the distances between the estimated trajectories and the GPS measurement at the resurfacing point ξ were calculated for the navigation solutions analysed. This value provides a measure of the drift of the solution, thus it was used as a term of comparison among the strategies considered. It is worth pointing out that navigation systems involving top notch sensors as DVL and FOG typically result in a drift of 2% – 5% of the total path length (Paull et al., 2014). However, ξ is not sufficient to evaluate the overall behaviour of the USBL-based method, which does not accumulate error over time, but it can still provide some useful insights on the local performance. Table 5 collects the results, for each mission of the dataset, concerning the presented metrics considered for the DVL and USBL sensors, respectively. It also reports the means of the metrics for each scenario, weighted according to the number of samples in each mission. Table 6, instead, gives the experimental results of the ξ distance at the resurfacing point for the three strategies presented on a mission-by-mission basis, as well as the percentage error on the total distance travelled. Lastly, in Fig. 9 the AUV trajectories obtained by the three navigation methods analysed within this work are depicted. Each figure represents the experimental results obtained from a representative mission randomly chosen for each scenario considered.

6.2. Seeps detection and geolocalisation

Regarding the gas seeps detection, Table 7 summarises the outputs of the ATR system. The algorithm was run onboard on FeelHippo AUV by exploiting the NVIDIA Jetson Nano. The collected results were subsequently analysed in a post-processing stage by a human operator that classified the outputs as True Positives (TP) and False Positives (FP), i.e. correct and wrong detections. This allowed to assess the

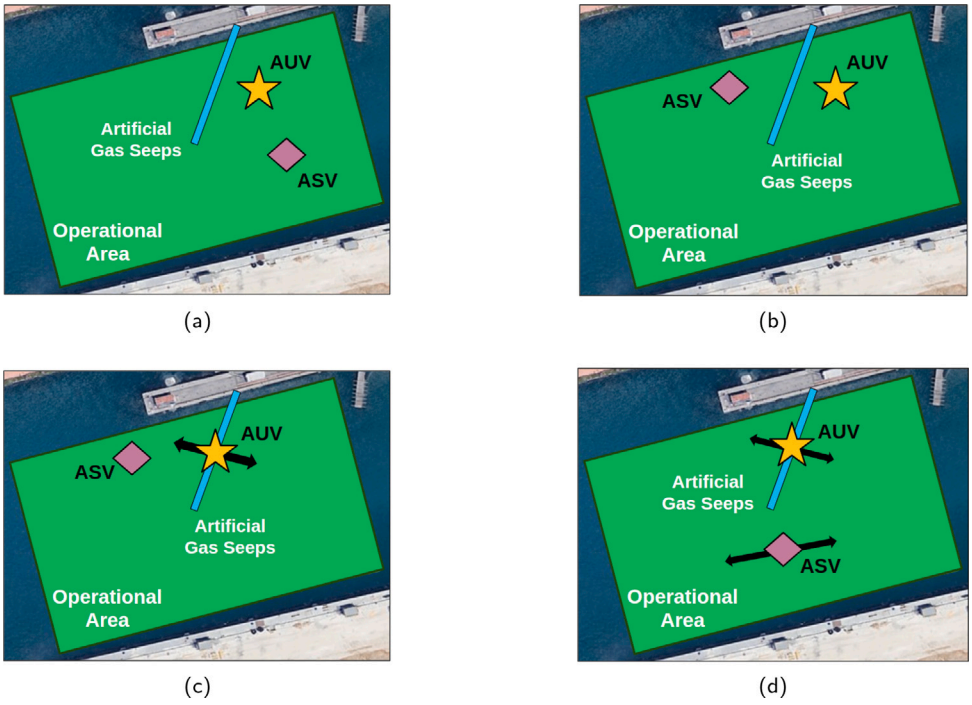


Fig. 8. Experimental scenario examples: (from left to right) in *S1* both vehicles operates in an area in the surrounding of the gas seeps; in *S2* the AUV moves far from the bubbles while the ASV is located at the opposite side of the pipe; *S3* involves the AUV moving on the seeps and ASV hovering in a fixed location at the opposite side of the pipe; finally in *S4* the AUV straddles the bubbles while the ASV moves at the centre of the basin.

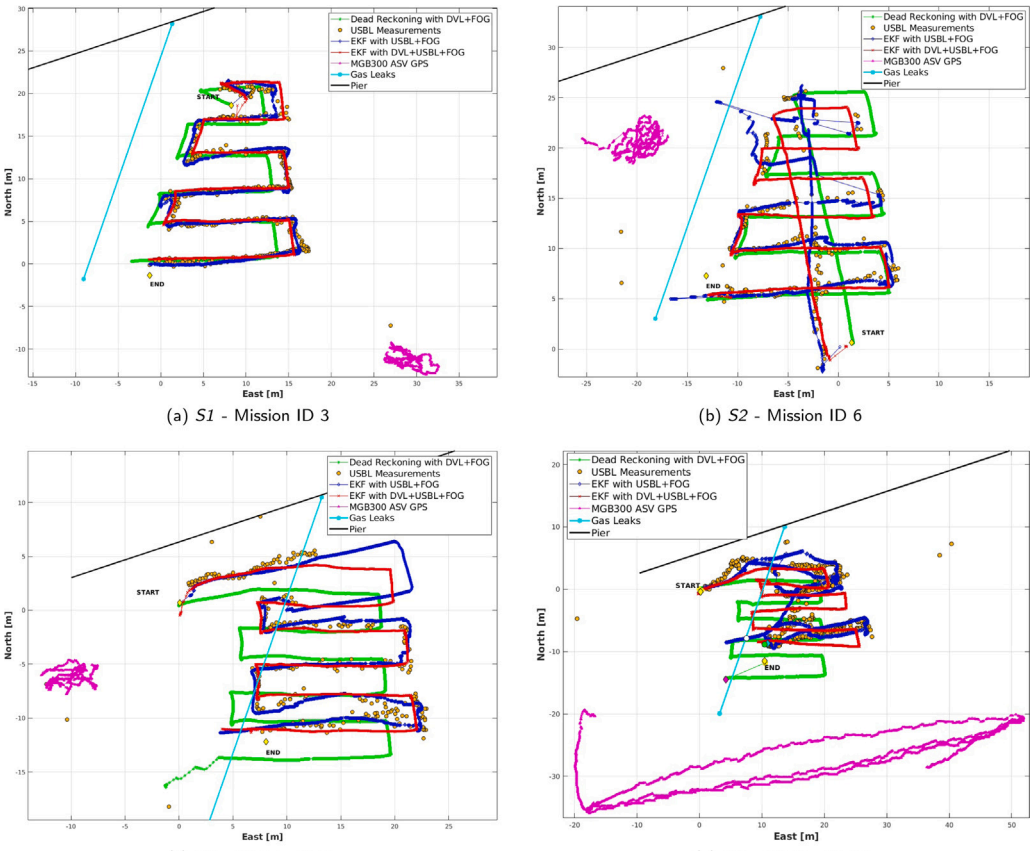


Fig. 9. North-East plots of the AUV trajectories, an example for each scenario. The plot shows the results of the three navigation solution proposed: DVL-based dead reckoning (in green), kinematic USBL-based EKF (in blue) and kinematic EKF exploiting both DVL and USBL observations (in red). Furthermore, the figures reports the USBL measurements (in orange), the position of the pipe generating the bubbles streams (in cyan), the ASV GPS data (in magenta) and the pier outline (in black). For each path, the first point represented corresponds to the instant at which the AUV reached the desired depth. These points, and the mission resurfacing points obtained from the GPS data, are represented as yellow rhombuses and labelled appropriately.

Table 5
Indicators of the effects of gas bubbles on DVL and USBL sensors for each mission.

ID	Scenario	DVL	USBL		
		Invalid [%]	l_p [%]	Δ_t [s]	$\overline{\chi}$ [Hz]
1		0.0	16.2	5.3	0.8
2		0.9	36.2	10.5	0.6
3	*S1*	0.0	18.3	7.8	0.8
4		0.0	20.2	8.0	0.8
Mean		0.2	22.2	7.7	0.8
5	*S2*	0.0	54.9	33.3	0.5
6		3.4	54.0	49.8	0.4
Mean		1.6	54.5	41.6	0.5
7		6.9	62.9	135.2	0.3
8	*S3*	8.4	43.9	114.9	0.5
9		13.5	34.2	29.8	0.6
Mean		9.5	46.8	92.6	0.5
10	*S4*	7.2	27.9	25.0	0.7

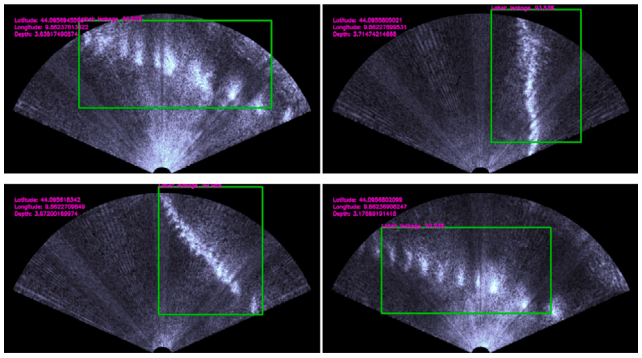


Fig. 10. Four detection examples performed online thanks to the developed ATR strategy with the SSD network during the experimental campaign conducted in April 2021 at the CSSN, La Spezia, Italy. In green the bounding box traced around the recognised seeps, while in purple, the localisation outcomes as geodetic coordinates.

performance of the developed ATR solution in terms of precision P, computed as:

$$P = \frac{TP}{TP + FP} \tag{11}$$

Four examples of TP detections are depicted in Fig. 10; while, for the sake of completeness, an example of FP is reported in Fig. 11.

The entire workflow of the ATR system, including the target geolocalisation, was validated during missions 2,3,5 and 6. This was due to the hardware problem encountered by the authors during the second day of activity (more details in Section 5.2) and to the fact that in mission 1 the gas pipeline was not enlighted with the FLS. The positions of only the ATR TPs, geolocated in real time using the combined approach, were then compared with the available reference on drilled pipe location. The results, in terms of mean and standard deviation on the distance between the detected seeps and the deployed pipe, are reported in Table 8 for each mission conducted during the first day of activity, and thus containing FLS data. Fig. 12 shows the geolocalisation results for Mission 6, where the ATR findings are represented as blue circles, the reference position of the pipe as a cyan line, and the estimated AUV trajectory obtained with the combined approach is coloured in red. Mission 6 was chosen as an example to represent the outputs of the proposed ATR strategy, due to its statistical significance for the large number of TPs it contains, but similar results emerged from the other available datasets.

7. Discussion

Since the percentage of invalid DVL readings, reported in Table 3, is zero (or close to zero) for all the missions in *S1*, it is reasonable

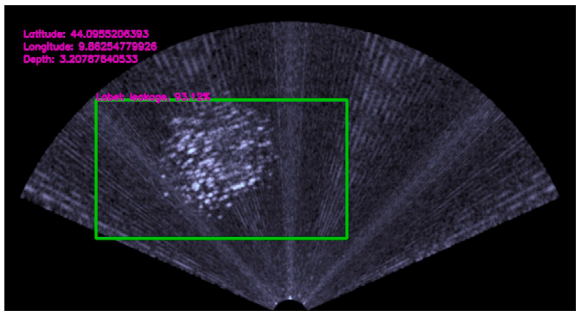


Fig. 11. An example of misidentified image where a possible shoal has been wrongly recognised as gas seep.

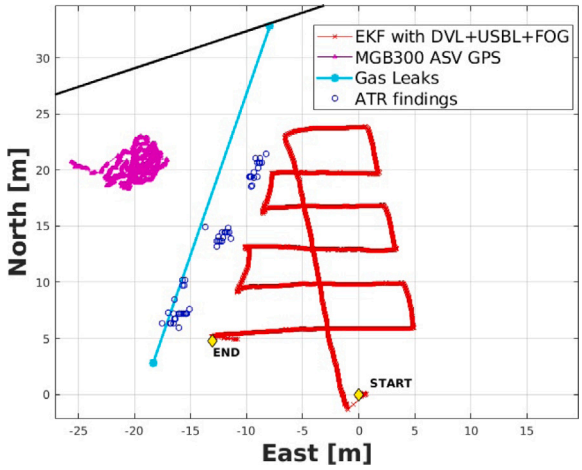


Fig. 12. The ATR findings (as blue circles) with respect the AUV trajectory estimated using the combined approach (in red) leveraging both DVL and USBL sensors. In cyan and black, the pipe from which the seeps were generated and the pier are shown, respectively. Finally, the position of the ASV is depicted in magenta.

to conclude that the DVL sensor is not influenced by the presence of gas bubbles in the nearby (tens of meters) of the AUV. Indeed, for such missions the drift of the navigation solution, measured by the ξ_{DVL} metric, is restricted within the 3.1% (in the worst case) of the distance travelled, in accordance with the typical results reported in literature (Paull et al., 2014). A similar result can also be found in *S2*, with a sole exception in mission 6, where the percentage of invalid DVL measures rises to 3.4%. However, the dead reckoning estimate leads to a drift of 2.4% of the total distance even for such mission. So, the percentage of invalid readings observed in these two scenarios (*S1* and *S2*) does not seem to significantly affect the performance of the DVL-based navigation strategy. Therefore, it is legitimate to regard the discussed results as benchmark performance for the DVL sensor and its tailored navigation solution. As consequence, the dead reckoning estimate can be considered reliable and used as reference trajectory while assessing the results for the USBL-based and combined approaches in scenarios *S1* and *S2*. A different situation arises in *S3* and *S4*, where the underwater robot moves above the pipe and the DVL readings are hence affected by the flow of gas bubbles. This was an expected outcome since the DVL sensor is mounted in a bottom looking configuration, so it should be affected by the presence of bubbles only when standing over them, as it happens in these scenarios. To support this conclusion, it can be seen that the average percentage of invalid measures is about $3-4$ times higher in *S3* and *S4* with respect to the worst case of mission 6 in *S2*. This is shown to have a negative impact on the accuracy of the vehicle trajectory estimation, with a significant increase in the drift, which reaches the 8.8% of the entire path length in the case of mission 8.

Table 6
Distance at the resurface point and percentage drift with respect to the travelled path of the examined navigation methods for each mission in the dataset.

ID	Scenario	ξ_{DVL} [m]	ξ_{DVL} [%PL]	ξ_{USBL} [m]	ξ_{USBL} [%PL]	$\xi_{Combined}$ [m]	$\xi_{Combined}$ [%PL]
1	*S1*	3.2	2.3	2.9	2.1	1.3	0.9
2		1.0	1.2	2.5	3.2	2.0	2.4
3		3.1	3.1	1.5	1.5	2.2	2.2
4		2.7	3.1	4.2	4.7	1.2	1.3
5	*S2*	4.1	3.6	2.2	1.9	0.7	0.6
6		2.3	2.4	2.4	2.3	2.0	2.1
7	*S3*	7.8	7.4	3.8	3.6	1.9	1.8
8		10.0	8.8	4.0	3.5	1.6	1.4
9		1.6	1.4	1.4	1.3	1.4	1.3
10	*S4*	7.1	6.5	4.4	4.0	2.5	2.3

Table 7
Results of the CNN-based ATR algorithm on FLS images.

ID	Scenario	TP	FP	Total	Precision
1	*S1*	–	–	–	–
2		37	8	45	0.82
3		23	5	28	0.82
5	*S2*	20	9	29	0.69
6		59	0	59	1.0

Table 8
Outcomes of the localisation strategy on the ATR TPs finding in terms of mean and standard deviation of the distance between the detected seeps and the deployed pipe.

ID	Scenario	Mean [m]	Standard deviation [m]
1	*S1*	–	–
2		3.33	1.33
3		0.99	0.66
5	*S2*	0.59	0.26
6		3.47	1.14

Table 9
Comparison between the dead reckoning estimated trajectory based on the DVL sensor and the USBL measurements.

	S1	*S2*
RMS error [m]	3.0	3.1
Max. error [m]	5.1	6.8
Min. error [m]	0.7	1.6
Mean error [m]	2.8	2.9
Std. dev. [m]	1.1	0.8

Regarding the USBL sensor, the analysed metrics highlight a notable deterioration of positioning system performance between missions performed in scenario *S1* and those conducted in scenarios *S2* and *S3* scenarios. Among the missions in *S1*, the USBL device provided a 22.2% of lost packets, a Δ_t of 7.7 s and a $\overline{\chi}$ of 0.8 Hz, in average. It is worth remember that the operational area represents a challenging scenario from the acoustic point of view (see 5.2 for more details), thus it is plausible to believe that such values represent the benchmark performance of the system in that area. The situation changes considerably in the scenarios *S2* and *S3*, in which the metric increases as gas bubbles interpose between the two vehicles, affecting positioning. In particular, the packet loss on average more than double, reaching the 54.5%; the Δ_t increased by 5 times up to an average of 41.6 [s]; while the update frequency is almost halved, reaching 0.45 [Hz] on average. To assess whether the impact of gas bubbles affected the USBL in terms of loss of communication or also had an influence on the accuracy of the sensor, the authors' computed the error between the dead reckoning estimated trajectory and the USBL measurements for the missions of the scenarios *S1* and *S2*. The analysis considered just these two scenarios because in both of them it was established by the previous results that the DVL-based solution can be considered as benchmark in a comparative study. The results, averaged among the missions of the same scenario, are collected in Table 9 and proved that the effect of the seeps on the USBL positioning system is limited to the loss of connection among the robots, thus to less measurements successfully delivered to the AUV, while no significant repercussion can be found on the measurements accuracy. In scenario *S4*, instead, an improvement in localisation performance can be appreciated compared with the *S2* and *S3* configurations. Although not as good as the benchmark case *S1*, the percentage of packets lost in *S4* is about half the mean value in *S2* and 0.6 times the mean in *S3*. $\overline{\chi}$ is also improved compared to the previous two scenarios, and as confirmed by the increase in the number of USBL measurements shown in Table 4, this means a more reliable connection between vehicles. In fact, it is reasonable to assume that the movement of the ASV in the middle of the basin allows to partially avoid the gas bubbles, resulting in better acoustic communication paths between vehicles.

The USBL-based method gives, in terms of distance ξ_{USBL} between the EKF estimate and the GPS data at the resurfacing point, in a limited error over all the dataset, with a maximum of 4.8% of the travelled distance in the worst-case scenario of mission 4. Nonetheless, this metric is not sufficient to conclude on the overall performance of the method, as it is an indicator of only the final portion of the AUV trajectory, where the effects of seepage are reduced and the acoustic connection between the two assets is more stable. In fact, as depicted by Fig. 9(b), estimation based solely on the USBL device is inaccurate in cases where no observations are available, resulting in jumps in the estimated trajectory; while it is more reliable in the area where USBL measurements are more frequent.

Finally, the combined approach provides a more robust and accurate estimate of vehicle position, avoiding the total lack of information due to the loss of connectivity of the USBL-based method and the ever-increasing drift typical of dead reckoning approaches. The results on $\xi_{Combined}$ prove that the approach performs globally better on every studied scenario, and the plots in Fig. 9 show that the estimate obtained with this method is generally more reliable and consistent compared to both the single sensor approaches.

Lastly, the developed ATR methodology allowed the underwater robot to detect and localise the simulated gas seeps online. Fig. 10 shows that the trained CNN architecture can identify the seeps in the FLS images even when enlighted from different views, proving its capability of generalising learned patterns. It is worth noting that due to the pipeline length utilised to simulate the gas seeps (about 35 m), in a single FLS frame only a portion of the seeps is captured. Additionally, the results reported in Table 7 demonstrate the recognition capability of the ATR solution. In fact, in both missions 2 and 3, a precision of 0.82 was achieved. By analysing the collected data, it emerged that the CNN misclassified some images where the FLS enlighted the quay. In mission 4 the precision decreases to 0.69. This is mainly due to the relatively high number of FP with respect to the TP caused by a school of fish, whose rendering in an FLS image could resemble a gas seep, as shown in Fig. 11. Then, in mission 6 an outstanding result was achieved

by having various TP without FP, helped by favourable environmental conditions.

The results regarding the online geolocalisation of the target represent a preliminary validation in an controlled scenario of the proposed algorithm. Although the geolocalisation accuracy is heavily dependent on the goodness of the navigation strategy, the outcomes reported above from the whole set of missions offer the opportunity to have a quantitative assessment of the correct performance of the overall system. More in detail, the geolocalisation accuracy has been evaluated in terms of the mean and standard deviation of the distance between the detected seeps and the deployed pipe. As can be seen from Table 8, the geolocalisation mean outcomes range from 0.59 to 3.47 m with, instead, the standard deviation metric assuming values ranging in the interval (0.26, 1.33) m. Finally, in order to provide metrics summarising the whole experimental activity, the results from the 4 considered missions have been averaged with mean and standard deviation of, respectively, 2.09 and 0.85 m.

8. Conclusions & future works

This manuscript addresses the problem of autonomously identify and localise underwater gas seeps in real-time using an AUV equipped with a FLS. To achieve this goal, it is necessary to be able to: (i) accurately estimate the absolute position of the vehicle in an environment with gas leaks presence; (ii) develop a solution to autonomously detect gas emissions in FLS imagery; (iii) retrieve the relative position of an object in the FLS acoustic image with respect to the sensor reference frame, and express this information as geodetic coordinates. Although the latter is an already solved problem (Negahdaripour, 2013), the former points have been scarcely addressed in literature, or approached by other means. To achieve this goal, the authors analysed the impact that gas emissions have on the acoustic sensors commonly exploited in AUV navigation strategies. The investigated technologies consisted of a DVL sensor and an USBL device, operating in the order of MHz and tens of kHz, respectively. For each technology, a tailored navigation solution was implemented and tested in at sea experiments, where an artificial gas source was reproduced. The results showed that both sensors suffer from disturbances caused by the presence of gas bubbles, which due to their strong scattering properties negatively affect the acoustic sensors functionality. Indeed, even though the studied sensors operate in different range of frequencies, it was proved that their performance are heavily influenced by gas seeps. The DVL sensor, entailing a dead reckoning technique, is more subject to invalid measurements and undetected outliers when operating above to the source of the gas emission. This leads to an increased drift in the position estimate which cannot be restored in the implemented navigation approach. The USBL device, instead, was mounted on an ASV to positioning the AUV and communicate the obtained data to the latter, exploiting a specifically designed acoustic protocol. The results showed that gas bubbles cause disturbances of the acoustic signals utilised by the device with a consequent inability to positioning or transmit the acquired information to the underwater vehicle, but no effects on the measurements accuracy were encountered. Furthermore, it was observed in Fig. 9 that connectivity losses between vehicles are more frequent near the pier, where gas flow is more intense, as it was reasonable to assume. The EKF, implemented to take advantage of the absolute position measurements of the AUV provided by the USBL, led to limited resurfacing errors, but lacked robustness and reliability when communication between the robots is lost. The method combining both sensors, on the other hand, proved to be more accurate and reliable across all the missions comprising the dataset, being robust to gas bubbles disturbances on the acoustic sensors, or at least mitigating their effects. Consequently, to improve the navigation results of an AUV moving in an area with gas seeps, it is recommended to use all available acoustic navigation sensors. The distinct frequency ranges and methods of operation enable differentiation in the nature of the information collected, which, when combined, allow to obtain a more accurate underwater navigation of the robot.

For what concerns the ATR of underwater gas seeps, instead, the authors proposed an online algorithm based on CNNs applied on FLS acoustic images. Once the targets were recognised, and their relative position with respect to the AUV computed, they were mapped by exploiting the absolute position of the underwater vehicle, estimated by its navigation system. The entire workflow of the algorithm was validated during at sea experiments, where all the detected targets were successfully geolocalised online, proving the effectiveness of the presented strategies. Moreover, the ATR method provided an average precision of about 0.86, in terms of the ratio between all TPs and the total number of targets detected in the five missions considered. The accuracy of the geolocalisation strictly depends on the AUV navigation accuracy. Here, the proposed EKF strategies can be pivotal, especially when operating in hazardous areas with gas leaks that could deteriorate DVL measurements. Finally, it is worth mentioning that the conducted experimental campaign showed that AI-based solutions can be utilised for detecting and localising gas leaks and CO_2 seeps in acoustic images, which is fundamental for maintenance and inspection surveys.

As far as the next stages are concerned, in order to improve and strengthen the ATR detection and localisation capability, classical computer vision strategies may be taken into account and added to the ATR loop as a final improving stage (for instance by detecting lighter pixels within the bounding box). Furthermore, the proposed solution will be tested in distinct artificial setups or even in a more realistic scenario which can indeed highlight whether such CNN-based architecture provides the user with adequate robustness and generalisation capability. More specifically, the developed strategy will be evaluated in a challenging scenario, such as the volcanic area involved in the DAMOSS1 trials. This will allow further validation of the results obtained in this work and will demonstrate the applicability of the ATR method to enable the execution of fully autonomous mission for detection and monitoring of natural gas seepage in real-world scenarios.

CRediT authorship contribution statement

Matteo Bresciani: Data curation, Investigation, Methodology, Software, Writing – original draft, Writing – review & editing. **Leonardo Zacchini:** Data curation, Investigation, Methodology, Software, Writing – original draft. **Alberto Topini:** Data curation, Methodology, Software, Writing – review & editing. **Alessandro Ridolfi:** Conceptualization, Funding acquisition, Project administration, Supervision. **Riccardo Costanzi:** Conceptualization, Funding acquisition, Project administration, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to thanks the CSSN members who provided precious support in carrying out the experiments at sea, perfectly incorporating the spirit of collaboration that is the SEALab laboratory's hallmark. Sincere thanks also go to Professor Francesco Maurelli from the Jacobs University of Bremen, Principal Investigator of the DAMOSS2 TNA within the European project EUMarineRobots, which received funding from the European Union's Horizon 2020 research and innovation program under grant agreement N° 731103.

References

Amran, I. Y., Isa, K., Kadir, H. A., Ambar, R., Ibrahim, N. S., Kadir, A. A. A., et al. (2021). Development of autonomous underwater vehicle for water quality measurement application. In *Proceedings of the 11th national technical seminar on unmanned system technology 2019* (pp. 139–161). Springer.

Antonelli, G., Indiveri, G., Barrera, C., Caccia, M., Dooly, G., Flavin, N., et al. (2021). Advancing the EU marine robotics research infrastructure network: the EU marine robots project. In *OCEANS 2021: San Diego – Porto* (pp. 1–10).

Asada, A., Ura, T., Maeda, F., Maki, T., Yamagata, Y., & Seiichi, T. (2010). Sub-bottom synthetic aperture imaging sonar system using an AUV and an autonomous surface tracking vehicle for searching for buried shells of toxic chemicals. In *2010 international waterside security conference* (pp. 1–3).

Bayat, B., Crasta, N., Crespi, A., Pascoal, A. M., & Ijspeert, A. (2017). Environmental monitoring using autonomous vehicles: a survey of recent searching techniques. *Current Opinion in Biotechnology*, *45*, 76–84.

Bhattarai, P., Krupiński, S., Unnithan, V., Maurelli, F., Secciani, N., Franchi, M., et al. (2021). A deep learning approach for underwater bubble detection. In *OCEANS 2021: San Diego – Porto* (pp. 1–5).

Blackford, J., Stahl, H., Bull, J., et al. (2014). Detection and impacts of leakage from sub-seafloor deep geological carbon dioxide storage. *Nature Climate Change*, *4*(11), 1011–1016.

Blomberg, A. E. A., Sæbø, T. O., Hansen, R. E., Pedersen, R. B., & Austeng, A. (2017). Automatic detection of marine gas seeps using an interferometric sidescan sonar. *IEEE Journal of Oceanic Engineering*, *42*(3), 590–602.

Bresciani, M., Peralta, G., Ruscio, F., Bazzarello, L., Caiti, A., & Costanzi, R. (2021). Cooperative ASV/AUV system exploiting active acoustic localization. In *2021 IEEE/RSJ international conference on intelligent robots and systems* (pp. 4337–4342).

Bresciani, M., Peralta, G., Ruscio, F., Costanzi, R., Zacchini, L., Franchi, M., et al. (2021). Localisation approaches for underwater autonomy within the euMarineRobots H2020 project: experimental activity at sealab. In *OCEANS 2021: San Diego – Porto* (pp. 1–5).

Bresciani, M., Tani, S., Ruscio, F., Zacchini, L., Bartalucci, L., Ridolfi, A., et al. (2022). Impact of natural gas seeps on the navigation of an autonomous underwater vehicle. In *2022 IEEE/OES autonomous underwater vehicles symposium* (pp. 1–6). IEEE.

Bucci, A., Zacchini, L., Franchi, M., Ridolfi, A., & Allotta, B. (2022). Comparison of feature detection and outlier removal strategies in a mono visual odometry algorithm for underwater navigation. *Applied Ocean Research*, *118*, Article 102961.

Carstensen, E. L., & Foldy, L. L. (1947). Propagation of sound through a liquid containing bubbles. *The Journal of the Acoustical Society of America*, *19*(3), 481–501.

Costanzi, R., Fanelli, F., Monni, N., Ridolfi, A., & Allotta, B. (2016). An attitude estimation algorithm for mobile robots under unknown magnetic disturbances. *IEEE/ASME Transactions on Mechatronics*, *21*(4), 1900–1911.

Costanzi, R., Fenucci, D., Caiti, A., Micheli, M., Vermeij, A., Tesei, A., et al. (2018). Estimation filtering for deep water navigation. *IFAC-PapersOnLine*, *51*(29), 299–304, 11th IFAC Conference on Control Applications in Marine Systems, Robotics, and Vehicles CAMS 2018.

De Beukelaer, S. M., MacDonald, I. R., Guinnasso, N. L., Jr., & Murray, J. (2003). Distinct side-scan sonar, RADARSAT SAR, and acoustic profiler signatures of gas and oil seeps on the gulf of Mexico slope. *Geo-Marine Letters*, *23*, 177–186.

Font, E. G., Bonin-Font, F., Negre, P.-L., Massot, M., & Oliver, G. (2017). USBL integration and assessment in a multisensor navigation approach for AUVs. *IFAC-PapersOnLine*, *50*(1), 7905–7910.

Fossen, T. I. (2002). Marine control systems–guidance. navigation, and control of ships, rigs and underwater vehicles. *Marine Cybernetics*, Trondheim, Norway, Org. Number NO 985 195 005 MVA, www. marinecybernetics. com, ISBN: 82 92356 00 2.

Franchi, M., Fanelli, F., Bianchi, M., Ridolfi, A., & Allotta, B. (2020). Underwater robotics competitions: The European robotics league emergency robots experience with FeelHippo AUV. *Frontiers in Robotics and AI*, *7*.

Franchi, M., Ridolfi, A., & Allotta, B. (2021). Underwater navigation with 2D forward looking SONAR: An adaptive unscented Kalman filter-based strategy for AUVs. *Journal of Field Robotics*, *38*(3), 355–385.

Huang, J., Rathod, V., Sun, C., Zhu, M., Korattikara, A., Fathi, A., et al. (2017). Speed/accuracy trade-offs for modern convolutional object detectors. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 7310–7311).

Hwang, J., Bose, N., Nguyen, H. D., & Williams, G. (2020). Acoustic search and detection of oil plumes using an autonomous underwater vehicle. *Journal of Marine Science and Engineering*, *8*(8).

Jin, L., Liang, H., & Yang, C. (2019). Accurate underwater ATR in forward-looking sonar imagery using deep convolutional neural networks. *IEEE Access*, *7*, 125522–125531.

Jones, D., Beaubien, S., Blackford, J., Foekema, E., Lions, J., De Vittor, C., et al. (2015). Developments since 2005 in understanding potential environmental impacts of CO2 leakage from geological storage. *International Journal of Greenhouse Gas Control*, *40*, 350–377.

Judd, A. G. (2003). The global importance and context of methane escape from the seabed. *Geo-Marine Letters*, *23*, 147–154.

Kato, N., Choyekh, M., Dewantara, R., Senga, H., Chiba, H., Kobayashi, E., et al. (2017). An autonomous underwater robot for tracking and monitoring of subsea plumes after oil spills and gas leaks from seafloor. *Journal of Loss Prevention in the Process Industries*, *50*, 386–396.

Kebkal, K., & Mashoshin, A. (2017). AUV acoustic positioning methods. *Gyroscopy and Navigation*, *8*, 80–89.

Kim, J., & Yu, S.-C. (2016). Convolutional neural network-based real-time ROV detection using forward-looking sonar image. In *2016 IEEE/OES autonomous underwater vehicles* (pp. 396–400). IEEE.

Klaucke, I., Weinrebe, W., Sahling, H., Bohrmann, G., & Börk, D. (2005). Mapping deep-water gas emissions with sidescan sonar. *Eos, Transactions American Geophysical Union*, *86*(38), 341–346.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, *25*.

Kvasić, I., Mišković, N., & Vukić, Z. (2019). Convolutional neural network architectures for sonar-based diver detection and tracking. In *OCEANS 2019 - Marseille* (pp. 1–6).

Leifer, I., Kamerling, M. J., Luyendyk, B. P., & Wilson, D. S. (2010). Geologic control of natural marine hydrocarbon seep emissions, coal oil point seep field, california. *Geo-Marine Letters*, *30*, 331–338.

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., et al. (2016). Ssd: Single shot multibox detector. In *European conference on computer vision* (pp. 21–37). Springer.

Liu, P., Wang, B., Deng, Z., & Fu, M. (2018). INS/DVL/PS tightly coupled underwater navigation method with limited DVL measurements. *IEEE Sensors Journal*, *18*(7), 2994–3002.

Lurton, X. (2002). *An introduction to underwater acoustics: principles and applications*. London: springer.

Maeda, Y., Shitashima, K., & Sakamoto, A. (2015). Mapping observations using AUV and numerical simulations of leaked CO2 diffusion in sub-seabed CO2 release experiment at ardmucknish bay. *International Journal of Greenhouse Gas Control*, *38*, 143–152, CCS and the Marine Environment.

Martin-Abadal, M., Guerrero-Font, E., Bonin-Font, F., & Gonzalez-Cid, Y. (2018). Deep semantic segmentation in an AUV for online posidonia oceanica meadows identification. *IEEE Access*, *6*, 60956–60967.

Maurelli, F., Krupiński, S., Xiang, X., & Petillot, Y. (2022). AUV localisation: a review of passive and active techniques. *International Journal of Intelligent Robotics and Applications*, *6*, 246—269.

Modasshir, M., Rahman, S., Youngquist, O., & Rekleitis, I. (2018). Coral identification and counting with an autonomous underwater vehicle. In *2018 IEEE international conference on robotics and biomimetics* (pp. 524–529).

Monk, S. A., Schaap, A., Hanz, R., Borisov, S. M., Loucaides, S., Arundell, M., et al. (2021). Detecting and mapping a CO2 plume with novel autonomous pH sensors on an underwater vehicle. *International Journal of Greenhouse Gas Control*, *112*, Article 103477.

Morgado, M., Oliveira, P., Silvestre, C., & Vasconcelos, J. (2007). Vehicle dynamics aiding technique for usbl/ins underwater navigation system. *IFAC Proceedings Volumes*, *40*(17), 111–116.

Nadimi, N., Javidan, R., & Layeghi, K. (2021). Efficient detection of underwater natural gas pipeline leak based on synthetic aperture sonar (SAS) systems. *Journal of Marine Science and Engineering*, *9*(11).

Negahdaripour, S. (2013). On 3-D motion estimation from feature tracks in 2-D FS sonar video. *IEEE Transactions on Robotics*, *29*(4), 1016–1030.

Nikolovska, A., Sahling, H., & Bohrmann, G. (2008). Hydroacoustic methodology for detection, localization, and quantification of gas bubbles rising from the seafloor at gas seeps from the eastern Black Sea. *Geochemistry, Geophysics, Geosystems*, *9*(10).

Nortek Manuals (2022). *Nortek manuals: DVL operations*. Norway: Nortek, Accessed: Jun. 28, 2022. URL https://support.nortekgroup.com/hc/en-us/articles/360029835811-Operations-Manual-DVL.

NVIDIA (2018). NVIDIA jetson nano. https://developer.nvidia.com/embedded/jetson-nano-developer-kit (last visit October 2021).

Palomeras, N., Furfaro, T., Williams, D. P., Carreras, M., & Dugelay, S. (2021). Automatic target recognition for mine countermeasure missions using forward-looking sonar data. *IEEE Journal of Oceanic Engineering*, *47*(1), 141–161.

Paull, L., Saeedi, S., Seto, M., & Li, H. (2014). AUV navigation and localization: A review. *IEEE Journal of Oceanic Engineering*, *39*(1), 131–149.

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 779–788).

Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. *Advances in Neural Information Processing Systems*, *28*.

Ridolfi, A., Costanzi, R., Fanelli, F., Monni, N., Allotta, B., Bianchi, S., et al. (2016). FeelHippo: A low-cost autonomous underwater vehicle for subsea monitoring and inspection. In *2016 IEEE 16th international conference on environment and electrical engineering* (pp. 1–6).

Rigby, P., Pizarro, O., & Williams, S. B. (2006). Towards geo-referenced AUV navigation through fusion of USBL and dvl measurements. In *OCEANS 2006* (pp. 1–6).

Rypkema, N. R., Fischell, E. M., & Schmidt, H. (2017). One-way travel-time inverted ultra-short baseline localization for low-cost autonomous underwater vehicles. In *2017 IEEE international conference on robotics and automation* (pp. 4920–4926). IEEE.

Schneider von Deimling, J., Brockhoff, J., & Greinert, J. (2007). Flare imaging with multibeam systems: Data processing for bubble detection at seeps. *Geochemistry, Geophysics, Geosystems*, *8*(6).

Taylor, P., Stahl, H., Vardy, M. E., Bull, J. M., Akhurst, M., Hauton, C., et al. (2015). A novel sub-seabed CO2 release experiment informing monitoring and impact assessment for geological carbon storage. *International Journal of Greenhouse Gas Control*, *38*, 3–17.

Thorsnes, T., Chand, S., Brunstad, H., Lepland, A., & Lågstad, P. (2019). Strategy for detection and high-resolution characterization of authigenic carbonate cold seep habitats using ships and autonomous underwater vehicles on glacially influenced terrain. *Frontiers in Marine Science*, *6*.

Urban, P., Köser, K., & Greinert, J. (2017). Processing of multibeam water column image data for automated bubble/seep detection and repeated mapping. *Limnology and Oceanography: Methods*, *15*(1), 1–21.

Valdenegro-Toro, M. (2016). Object recognition in forward-looking sonar images with convolutional neural networks. In *OCEANS 2016 MTS/IEEE monterey* (pp. 1–6). IEEE.

Weber, T. C., Jerram, K., & Mayer, L. (2012). Acoustic sensing of gas seeps in the deep ocean with split-beam echosounders. *Vol. 17*, In *Proceedings of meetings on acoustics ECUA2012*. Acoustical Society of America.

Weiss, K., Khoshgoftaar, T. M., & Wang, D. (2016). A survey of transfer learning. *Journal of Big data*, *3*(1), 1–40.

Zacchini, L., Franchi, M., Manzari, V., Pagliai, M., Secciani, N., Topini, A., et al. (2020). Forward-looking sonar CNN-based automatic target recognition: an experimental campaign with FeelHippo AUV. In *2020 IEEE/OES autonomous underwater vehicles symposium* (pp. 1–6). IEEE.

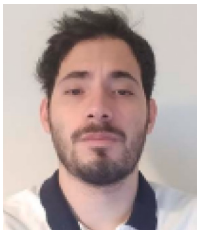
Zacchini, L., Ridolfi, A., Topini, A., Secciani, N., Bucci, A., Topini, E., et al. (2020). Deep learning for on-board auv automatic target recognition for optical and acoustic imagery. *IFAC-PapersOnLine*, *53*(2), 14589–14594.

Zacchini, L., Topini, A., Franchi, M., Secciani, N., Manzari, V., Bazzarello, L., et al. (2022). Autonomous underwater environment perceiving and modeling: An experimental campaign with feelhippo auv for forward looking sonar-based automatic target recognition and data association. *IEEE Journal of Oceanic Engineering*, *48*(2), 277–296.

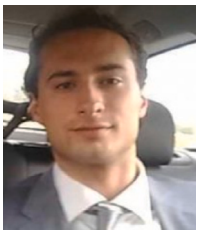
Zeiler, M. D., & Fergus, R. (2014). Visualizing and understanding convolutional networks. In *European conference on computer vision* (pp. 818–833). Springer.

Matteo Bresciani received the M.Sc. degree with honors in Robotics and Automation Engineering and the PhD from the University of Pisa, Italy, in December 2018 and May 2023, respectively. His PhD dissertation focused on cooperative solutions to improve the autonomy of marine vehicle, and he is now a Postdoctoral Researcher at the Department of Information Engineering, University of Pisa. His research interests include navigation and localisation of underwater robots, cooperative systems and optimal motion planning.

Leonardo Zacchini received the M.S. degree in Automation and Control Engineering, and the Ph.D. degree with a thesis focused on Autonomous Inspection Strategies for Underwater Robots from the University of Florence, Italy, in 2018 and 2021, respectively. He is currently a Postdoctoral Researcher in Robotics at the University of Florence, Italy. His research interests include guidance, navigation, and control systems for mobile robots, underwater robotics, robotics exploration, motion planning, and AI for robotics.



Alberto Topini received the M.Sc. in Artificial Intelligence from the University of Edinburgh, Edinburgh, Scotland, UK, in November 2019 and the Ph.D degree in Industrial Engineering from the University of Florence, Florence, Italy, in June 2023. He is currently carrying out research as Postdoctoral Researcher at the Department of Industrial Engineering of the University of Florence. His current research interests include decision-making, AI planning, probabilistic world modelling, and deep learning-based automatic target recognition methods for underwater robotics.



Alessandro Ridolfi received the Ph.D. degree in industrial engineering from the University of Florence, Florence, Italy, in 2014. He is currently a Ph.D. Researcher and an Assistant Professor of machine theory and robotics with the School of Engineering, Department of Industrial Engineering, University of Florence, Italy. His current research interests include biorobotics, vehicle dynamics, mechanical systems modelling, robotics, and underwater robotics.

Riccardo Costanzi received the Ph.D. degree in industrial engineering from the University of Florence, Florence, Italy, in 2015. He is currently an Associate Professor in Automatic Control at the School of Engineering, University of Pisa, Pisa, Italy. His research activity is focused on underwater robotics with particular interest to navigation and control systems for autonomous underwater vehicles.